Niveau(x) d'enseignement: 4 Numéro EDUQ : 9170978
Centre d'intérêt: 841

Test de mathématique : niveau secondaire : guide d'interprétation / Direction de la mesure et de l'évaluation des apprentissages. -- [Québec] : Direction générale du développement pédagogique, [1979?].
38 p. : ill. ; 28 cm.

Titre de la couv. Code: 16-7534.

Vedette(s)-secondaire(s): Québec (Province). Ministère de l'éducation. Direction de la mesure et de l'évaluation des apprentissages • Québec (Province). Direction générale du développement pédagogique

Descripteur(s): Mathématique • Rendement scolaire • Note • Interprétation normative • Secondaire (Ordre d'enseignement)

Identificateur(s) de sujet: Test standardisé

Ce document vient compléter la série de manuels ou guides d'accompagnement du test de mathématique (240 307) offert par le ministère de l'Education aux élèves de quatrième secondaire. A l'intention des professionnels enseignants, et même non enseignants, aussi bien que des administrateurs scolaires, ce document tente de synthétiser les renseignements nécessaires pour une utilisation aussi judicieuse que possible des résultats obtenus à ce test.

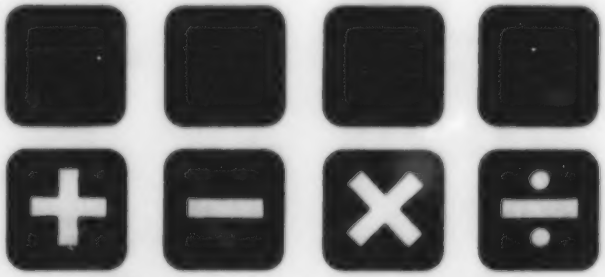
Type de document:

Gouvernement du Québec
Ministère de l'Éducation
Direction générale
du développement pédagogique

CADRE
CENTRE DE DOCUMENTATION
1940, BOUL. HENRI-BOURASSA EST
MONTRÉAL H2B 1S2

9170978

guide d'interprétation



test

DE MATHÉMATIQUE

niveau secondaire

Tous droits de reproduction, de traduction et d'adaptation réservés au ministère de l'Éducation

Direction de la mesure et de l'évaluation des apprentissages
7979

18-7634

CADRE

Tous droits de reproduction, de traduction et d'adaptation
réservés au ministère de l'Education

Dépôt légal - 3e trimestre 1979
Bibliothèque nationale du Québec

COMITÉ RESPONSABLE DE L'ÉLABORATION DU TEST DE MATHÉMATIQUE

La responsabilité de superviser et de coordonner toutes les étapes du développement du test a été confiée à un comité formé de spécialistes en mathématique et en mesure et évaluation.

Ce comité a tenu plusieurs réunions pour établir les objectifs et le contenu du test, suivre les différentes étapes de son expérimentation et finalement décider des modalités de son application. Toute décision concernant le contenu du test a été strictement la responsabilité du comité.

Nous remercions particulièrement les membres suivants, qui en ont fait partie au cours de l'année 1978-79. Ils ont permis, par leur travail assidu, que le test réponde mieux aux fins proposées:

Membres du comité (78-79)

Hélène Bouchard, conseiller en mathématique, C.S.R. Jean-Talon;

Réal Béland, responsable des programmes de mathématique au secondaire, D.G.D.P. (MEQ);

Jean-Marcel Bourgault, conseiller en mesure et évaluation, D.G.D.P. (MEQ);

Pierre Lafleur, conseiller en mesure et évaluation, C.S.R. Blainville - Deux-Montagnes;

Michel Labelle, conseiller en mesure et évaluation, C.S. de Huntingdon;

Granville E. Miller, Pedagogical Consellor, South Shore R.S.B.

Jean Pierre, professeur de mathématique, C.S. Chomedey-de-Laval;

Hélène Tessier, conseiller en mathématique, C.S. Baldwin-Cartier.

TABLE DES MATIÈRES

PAGES

INTRODUCTION..... 5

1. PRÉSENTATION DES RÉSULTATS..... 6

1.1 Méthode de correction..... 7

1.2 Publication des résultats..... 9

2. INTERPRÉTATION DES RÉSULTATS..... 16

2.1 Statistiques d'ensemble..... 16

2.2 Indices de réussite et de discrimination..... 18

2.3 Répartition des résultats..... 24

2.4 Liste des normes..... 26

3. CONSIDÉRATIONS TECHNIQUES..... 31

3.1 Validité du test..... 31

3.2 Fidélité du test..... 34

CONCLUSION: Qui interprète les résultats des tests?..... 38

LISTE DES TABLEAUX ET DES FIGURES

		PAGE
Figure 1	Equivalence des différentes échelles de scores utilisées	
Exemple 1	Résultats (individuels) de tests	
Exemple 2	Liste des résultats	
Exemple 3	Statistiques d'ensemble	
Exemple 4	Indices de réussite	
Exemple 5	Répartition des résultats individuels	
Exemple 6	Liste des normes	
Tableau 1A	Statistiques d'ensemble pour le Québec (population totale)	
Tableaux 1B et 1C	Statistiques d'ensemble pour les voies 281-422 (régulière) et 281-432 (enrichie)	
Tableau 2	Indices de réussite et de discrimination (Québec)	
Figures 2 à 6	Répartition des résultats: test et sous-tests (Québec)	
Tableau 3A	Normes - ensemble du Québec, résultat global	
Tableau 3B	Normes - ensemble du Québec, sous-tests	
Tableau 4	Normes - ensemble du Québec, voies 281-422 et 281-432	
Tableau 5	Corrélations - test et sous-tests	
Tableau 6	Consistance interne et autres données relatives	
Tableau 7	Corrélations test vs certains examens de rendement scolaire	

INTRODUCTION

Le présent document vient compléter la série de manuels ou guides d'accompagnement du test de mathématique (240 307) offert par le ministère de l'Education aux élèves de quatrième secondaire.

A l'intention des professionnels enseignants, et même non enseignants, aussi bien que des administrateurs scolaires, ce document tente de synthétiser les renseignements nécessaires pour une utilisation aussi judicieuse que possible des résultats obtenus à ce test.

Dans toutes les phases de son élaboration, comme le définit le Règlement no 7, on a voulu que ce test soit un "instrument standardisé préparé de façon scientifique". Dans l'interprétation des résultats, les utilisateurs auront ainsi profit à en connaître toutes les dimensions pour en apprécier les capacités et les limites. On ferait bien, dans cette opération, de garder continuellement à l'esprit qu'il implique les considérations suivantes:

1. Si les intentions et le contenu du test doivent être confrontés avec le contenu et les objectifs des programmes, un effort d'épuration, par l'expérimentation et l'analyse d'item, a permis l'élimination des questions dont la difficulté et le pouvoir de discrimination sont douteux;

2. Bien qu'une tentative ait été faite pour standardiser l'administration et la correction du test, les circonstances particulières qui prévalent dans la classe au moment même de l'administration ne sont pas sans influencer les résultats individuels;

3. Enfin une étude de la validité et de la fidélité du test a été entreprise, de même que la transformation des résultats bruts en scores plus significatifs applicables à un grand nombre de candidats (établissement de normes), dans le but d'arriver à une interprétation plus objective des résultats en fonction des intentions du test.

1. PRÉSENTATION DES RÉSULTATS

Le programme de tests du M.E.Q. se propose de fournir des indices d'ordre général sur la qualité des apprentissages scolaires mesurés (1):

Les tests veulent contribuer, par l'identification des forces et des faiblesses des élèves, à la solution des problèmes pédagogiques dans une perspective dynamique. Face à une situation globale, on ne veut pas mettre en cause l'un plus que l'autre les différents facteurs qui peuvent intervenir dans l'apprentissage. On veut surtout éviter de donner l'impression qu'on cherche, par ce moyen, à évaluer le rendement des enseignants.

Notre but n'est pas de rappeler ici tous les objectifs du test de mathématique offert en quatrième secondaire (240 307). On pourra en trouver une description détaillée dans le "manuel d'information" qui accompagne le test (2). Qu'il nous suffise de rappeler ici que celui-ci veut répondre aux objectifs généraux du programme de tests du M.E.Q.

A cette fin, ce test s'efforce de mesurer un certain nombre d'objectifs poursuivis dans l'apprentissage de cette matière au cours des quatre premières années du secondaire. Ces objectifs que l'on peut considérer comme prioritaires constituent un certain éclairage sur les éléments mesurables dans cet apprentissage.

Cependant un instrument comme celui-ci n'est pas conçu pour mesurer l'atteinte ou non de tous les objectifs d'un programme particulier. Pour cette raison, et pour d'autres, il ne vise pas la sanction des études à des fins de promotion ou de classement. Il n'en a ni l'intention ni la portée.

---

(1) Cf. Programme de base de tests à partir de l'année 1976-77, dans Recueil des règles de gestion des commissions scolaires, document 12-04-02, mars 76.

(2) M.E.Q. Manuel d'information, Test de mathématique, niveau secondaire, 1978-79, pp. 4 à 10.

1.1 Méthode de correction

La correction des feuilles de réponses, effectuée par le MEQ, est informatisée et sert à établir les résultats bruts ainsi que les résultats normalisés.

1.1.1 Les résultats bruts

Les résultats bruts sont simplement l'expression du nombre de "bonnes réponses". Ainsi chaque bonne réponse compte pour un (1) point:

(1) Une bonne réponse est celle où le choix que l'on considère correct a été indiqué et seulement celui-là;

(2) Une réponse omise est considérée comme une mauvaise réponse.

1.1.2 Les résultats standardisés et normalisés

Les résultats bruts sont transformés en scores standardisés ou normalisés qui expriment directement la valeur de la performance fournie par rapport à celle de la population testée: du Québec, de l'organisme, de l'école. Les scores sont ainsi présentés dans différentes échelles, selon la nature et la précision des renseignements désirés.

(A) L'échelle des cotes standardisées (1)

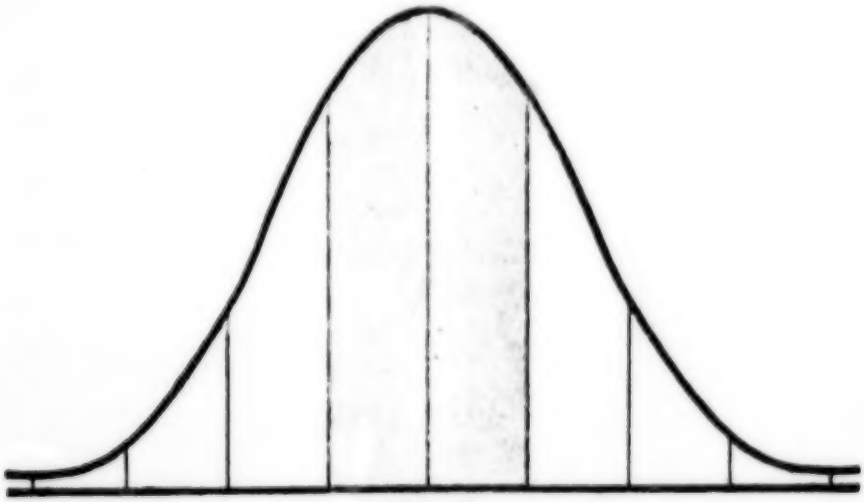
A l'intention des parents surtout, on a voulu exprimer les résultats à partir d'une échelle qui ne nécessite pas de longues explications et qui puisse facilement être interprétée, à partir d'un ordre de grandeur connu.

Pour éviter de véhiculer la notion d'échec en communiquant les résultats des tests, et compte tenu d'un niveau de difficulté variable pour chacun d'eux, on a considéré qu'une moyenne arbitraire de 63 serait plus valable. C'est pourquoi les résultats ont été transformés en cotes standardisées à partir de cette moyenne de 63, qui a été attribuée à l'ensemble de la population du Québec soit pour l'ensemble du test, soit pour l'un ou l'autre de ses sous-tests. L'écart-type dans cette distribution a été fixé à 10. Les moyennes locales varient en fonction de cette moyenne du Québec, selon les résultats obtenus au niveau des organismes.

---

(1) Voir Figure 1, à page suivante.

Figure 1 - Relations entre différentes échelles utilisées et interprétation des cotes



(1) Cette échelle des scores standardisés ne s'applique que dans le cas de la distribution normale. Dans le cas des tests du MEQ, il faut s'attendre à une certaine variabilité, selon que les résultats se rapprochent plus ou moins d'une distribution normale.

(B) L'échelle des stanines

Elle regroupe les résultats en neuf catégories, dans une distribution normalisée. Le stanine 5 est attribué à 20% des individus qui ont obtenu un résultat moyen, et les autres cotes sont attribuées par référence à celle-ci.

Ainsi, le stanine 1 est attribué aux individus donr les résultats sont parmi les 4% inférieurs, tandis que le stanine 9 qualifie les résultats qui se situent dans les 4% supérieurs.

(C) L'échelle des centiles

L'échelle des centiles indique le rang des sujets par rapport au pourcentage des individus qui ont obtenu une note inférieure à la leur: en d'autres termes, elle va de 1 à 99 et le rang centile attribué indique le pourcentage des individus qui ont eu un résultat inférieur à ce rang centile. Ainsi, le rang centile 62 indique que le ou les candidats auxquels ce rang centile est attribué ont obtenu un résultat supérieur à 62% de ceux qui ont subi le même test.

1.2 La publication des résultats

A la suite de l'administration du test, le MEQ transmet aux organismes les trois (3) sortes de rapports décrits ci-dessous.

1.2.1 Les "Résultats de tests" destinés aux élèves ou aux parents

Ces relevés fournissent aux candidats ou aux parents les résultats (RES) que les candidats ont obtenus au test entier seulement. Ils donnent également la moyenne locale (commissions scolaires et institutions privées comprises, s'il y a lieu) et celle du Québec, afin de situer les résultats des candidats dans les diverses populations qui ont passé le test.

Sur ces relevés, tous les résultats ne sont exprimés qu'en cotes standardisées.

Ces relevés sont adressés à la commission scolaire ou à l'institution pour être remis aux élèves ou aux parents.

Gouvernement du Québec
Ministère de l'Éducation
Direction générale
du développement pédagogique

Résultats de tests

000018 TESTS RENO [illegible] - 1979 790316

MASL - 1503-6308

[illegible]

[illegible]

146307 Mathématique - global	74 e	64 83

[illegible]

1.2.2 La liste des résultats des candidats destinée aux écoles (Exemple 2 ci-dessous).

Cette liste, adressée aux organismes, renferme les résultats individuels. En plus des résultats contenus sur le relevé de notes décrit ci-dessus, elle contient les résultats bruts (RES) obtenus au test et aux sous-tests.

Elle renferme de plus les résultats bruts maximums (RES MAX); le stanine pour le Québec (STA QUE); le centile pour la commission scolaire et le Québec; le score standardisé pour le Québec (S.ST QUE); les moyennes pour la commission scolaire, la région et le Québec. Tous ces résultats sont exprimés à la fois pour le test entier et chacun des sous-tests.

EXEMPLE 2

TRAITEMENT INFORMATISE DE TESTS

LISTE DES RESULTATS

013 CHAMBLY (DE)
004 ECOLE SECONDAIRE CATHOLIQUE

TEST DE MATH AU SEC

CODES PERMAN. ET TEMPORAIRE	NOM ET PRENOM DU CANDIDAT ADRESSE DU CANDIDAT	NUMERO ET NOM DU TEST OU DU SOUS-TEST	RES MAX	RES QUE	STA QUE	CENTILE CS	CENTILE QUE	S.ST QUE	MOY/ECART CS	MOY/ECART REG	MOY/ECART QUE
BERM-1951-GEOU BERTHELOT 013-004-004		240307001 ENSEMBLES-TEST DE NOMBRES	5	10	3	25	22	57	6	6	6
		240307002 CALCUL ALGEBR.-EQUATIONS	8	15	5	51	42	62	7	6	8
		240307003 GEOMETRIE	8	10	6	71	70	71	6	6	5
		240307004 FONCTIONS-PLAN CARTESIEN	11	15	6	88	77	73	7	6	6
		240307 MATHEMATIQUE GLOBAL	34	50	6	80	74	70	25	21	23

Ainsi le candidat ci-dessus, au premier sous-test, a obtenu 5 sur 10 en scores bruts. C'est ce qui le situe au 3e stanine en rapport à la population du Québec, au 25e et 22e centile respectivement en rapport aux populations de sa commission scolaire et du Québec. En score standardisé, à ce sous-test, il est au-dessous de la moyenne du Québec (63) puisqu'il a 57. Les moyennes en scores bruts sont respectivement de 6,4 et 6 pour sa commission scolaire, sa région administrative et le Québec.

1.2.3 Diverses données qui décrivent les résultats d'ensemble des commissions scolaires ou des institutions privées

Ces données comprennent pour l'organisme et chacune des écoles, où le nombre d'élèves est supérieur à 19:

(A) Un tableau de "statistiques d'ensemble":

Ce tableau décrit, par diverses statistiques, la tendance centrale (moyenne, médiane, mode) ainsi que la variabilité (écart-type, étendue simple) des résultats.

Voici à titre d'exemple, pour une commission scolaire.

EXEMPLE 3

TRAITEMENT INFORMATISE DE TESTS

S T A T I S T I Q U E S D ' E N S E M B L E

NUMERO ET NOM DE L'OPERATION 740000 TEST DE MATH 40 SEC ANGL L ENS
NUMEROS DE DEMANDES D'INSCRIPTION ET DE CHEQUETTE 001 ET 781
NUMERO ET TITRE DE DEMANDE DE CALCUL DE NORMES 782 NORMES OPERATION 740000
NUMERO ET TITRE DE DEMANDE D'EXTRACTION 782 MATH TEST SECONDARY 4

**************** 630000 SOUTH SHORE (PROTESTANTS)

	MOYENNE	MAXIMUM	ECART TYPE	MODE	MEDIANE	ETENDUE SIMPLE	NOMBRE DE CANDIDATS
740307001 ENSEMBLES SYST DE NOMBRES	005,26	10	002,06	007	006,77	10	628
740307002 CALCUL ALGEBRIQUE EQUATIONS	009,86	15	003,23	009	009,67	15	628
740307003 GEOMETRIE	006,82	10	002,12	007	005,58	10	628
740307004 FONCTIONS PLAN CARTESIEN	007,73	15	305,36	006	006,32	15	628
740307 MATHEMATIQUE	030,19	50	008,91	028	027,40	45	628

Le tableau comporte les principaux renseignements suivants:

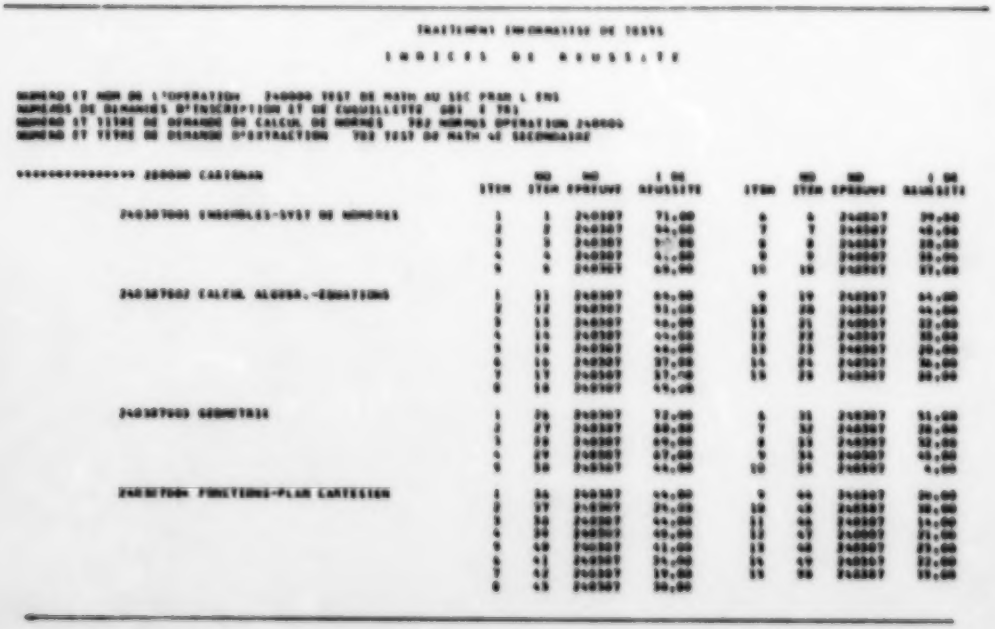
Maximum	Résultat maximum que les candidats pouvaient obtenir (nombre d'items).
Ecart-type	Mesure de la dispersion ou de la variabilité des scores.
Mode	Résultat qui se rencontre le plus fréquemment dans la série des scores.

Médiane	Point sur l'échelle des scores au-dessous ou au-dessus duquel il y a 50% des scores (sujets). Point milieu de la distribution des scores.
Etendue simple (totale)	Distribution des scores obtenus. Mesure de variabilité établie à partir du score le plus bas et du score le plus élevé.

(B) Un tableau des "indices de réussite":

Celui-ci indique en donnant le nombre d'items (1ère col.) et le numéro de l'item dans le test (2e col.), le pourcentage de succès à chacun, pour la population de l'école ou de la commission scolaire.

EXEMPLE 4

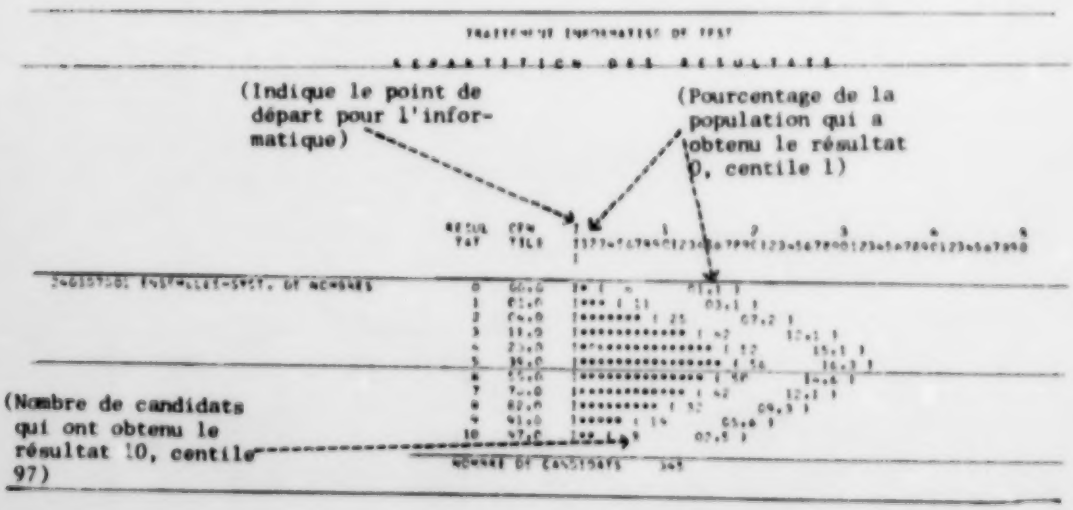


(C) La "répartition des résultats" au test et aux sous-tests:

Ce tableau indique où chaque résultat (1re colonne) se situe dans l'échelle des centiles (2e colonne). Un histogramme de fréquence (formé par des astérisques) fait la distribution des scores en pourcentage (1re ligne du tableau: de 1% à 50%); pour simplifier la lecture de l'histogramme, entre parenthèses, on indique le nombre et le pourcentage de candidats qui ont obtenu le score correspondant.

Voici, à titre d'exemple, l'un de ces tableaux, pour une commission scolaire, et l'un des sous-tests.

EXEMPLE 5



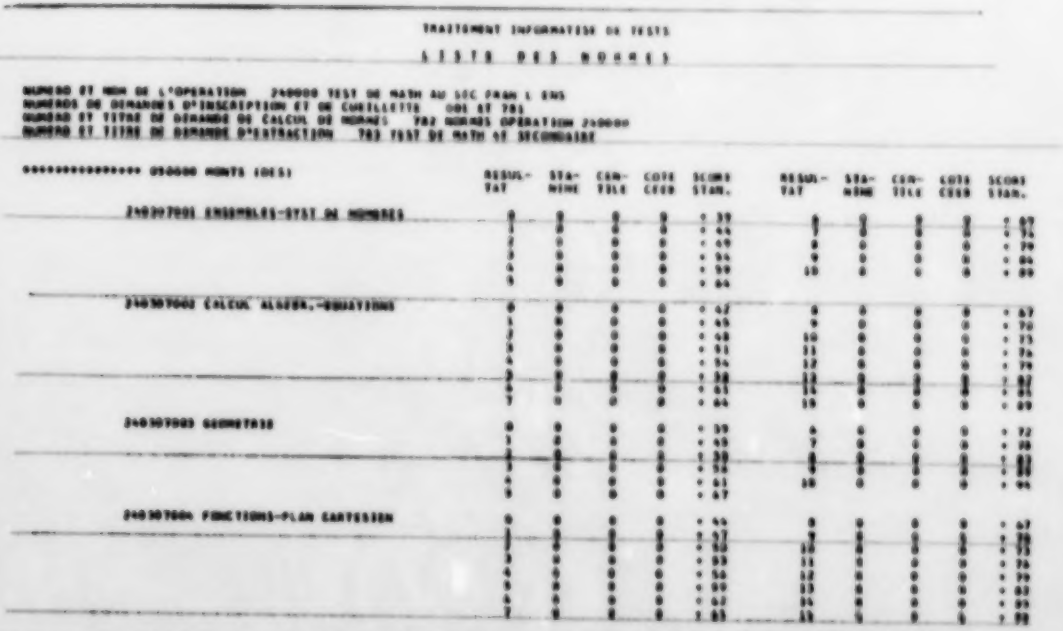
(7) La liste des normes établies pour la commission scolaire ou l'école

Cette liste donne les "normes locales". Elle permet de dire où se situe le résultat brut d'un élève par rapport à celui des autres élèves de sa commission scolaire ou de l'école. Les résultats bruts obtenus au test ou aux sous-tests sont ainsi transformés en cotes standardisées, en rangs centiles ou en stanines, par rapport à la population des organismes ou des écoles.

Les normes établies pour l'ensemble d'une commission scolaire comprennent les données des institutions privées situées dans leur territoire.

Il est évident que les données (stanines et centiles) établies à partir des résultats d'une petite population ne peuvent être vraiment significatifs.

EXEMPLE 6



N.B.: Dans l'exemple ci-dessus, le stanine, le centile et la cote CEEB n'ont pas été calculés à cause d'un trop petit nombre d'élèves.

2. INTERPRÉTATION DES RÉSULTATS

On trouvera ci-dessous les données obtenues par l'analyse des résultats pour l'ensemble de la population (23 913) qui a passé le test au Québec en avril 1979.

Ces données permettront d'établir des comparaisons avec les résultats obtenus au niveau local.

Afin de faciliter l'étude, nous adopterons, dans la présentation, l'ordre suivi par le service de l'Informatique du MEQ (SIMEQ) dans la publication des résultats adressés aux organismes et aux écoles.

2.1 Statistiques d'ensemble pour le Québec

Le tableau 1 présente des statistiques d'ensemble, à partir des scores bruts, pour tous les candidats (23 913) qui ont passé le test dans l'ensemble du Québec en avril 79. Il permettra des comparaisons avec les statistiques similaires publiées pour une école ou une commission scolaire.

TABLEAU 1 A - **Test de mathématique 240 307**

Mesures de tendance centrale et de variabilité des scores

Ensemble de la population (23 913) qui a passé le test au Québec, avril 1979.

TRAITEMENT INFORMATISE DE TESTS

S T A T I S T I Q U E S D ' E N S E M B L E

NUMERO ET NOM DE L'OPERATION 240000 TEST DE MATH AU SEC SECTEUR FR
NUMEROS DE DEMANDES D'INSCRIPTION ET DE CUEILLETTE 001 ET 001
NUMERO ET TITRE DE DEMANDE DE CALCUL DE NORMES 001 NORMES OPERATION 240000 1979
NUMERO ET TITRE DE DEMANDE D'EXTRACTION 000 AUCUNE EXTRACTION

**************** GROUPE NIVEAU PROVINCIAL

		MOYENNE	MAXIMUM	ECART-TYPE	MODE	MEDIANE	ETENDUE SIMPLE	NOMBRE DE CANDIDATS
240307001	ENSEMBLES SYST DE NOMBRES	008,59	10	002,22	009	009,10	10	23913
240307002	CALCUL ALGEBRIQUE EQUATIONS	008,19	15	003,56	007	007,85	15	23913
240307003	GEOMETRIE	008,17	10	002,12	006	006,60	10	23913
240307004	FONCTIONS PLAN CARTESIEN	006,67	15	003,38	006	006,75	15	23913
240307	MATHEMATIQUE	025,63	50	009,07	024	025,26	48	23913

Les tableaux 1A et 1B présentent les mêmes statistiques que le tableau 1, cependant en tenant compte de l'inscription des étudiants en mathématique, soit les voies 281-422 et 281-432 respectivement. Ils sont établis à partir d'un échantillon de 2500 candidats, pour les écoles qui ont nettement identifié leur population inscrite au test. On devra les interpréter en conséquence.

TABLEAU 1 B - **Test de mathématique 240 307**

Mesures de tendance centrale et de variabilité des scores - Inscription à la voie 281-422, pour un échantillon de 2 500 candidats au hasard de la population du Québec.

STATISTIQUES D'ENSEMBLE

Niveau provincial 000 000

782 - (281-422-Maths) - Voie régulière

		MOYENNE	MAXIMUM	ÉCART-TYPE	MODE	MÉDIANE	ÉTENDUE SIMPLE
240307001	Éléments de mathématiques	6,81	10	1,81	005	6,57	10
240307002	Calcul algébrique	8,67	15	2,55	006	5,86	15
240307003	Géométrie	6,48	10	1,80	006	6,00	10
240307004	Plan cartésien	5,58	15	2,52	005	4,89	15
240307	Mathématique globale	21,02	50	6,18	020	19,92	43

TABLEAU 1 C - **Test de mathématique 240 307**

Mesures de tendance centrale et de variabilité des scores - Inscription à la voie 281-432, pour un échantillon de 2 500 candidats au hasard de la population du Québec.

STATISTIQUES D'ENSEMBLE

Niveau provincial 000 000

783 - (281-432-Maths) - Voie enrichie

		MOYENNE	MAXIMUM	ÉCART-TYPE	MODE	MÉDIANE	ÉTENDUE SIMPLE
240307001	Éléments de mathématiques	6,74	10	1,68	007	6,60	10
240307002	Calcul algébrique	10,25	15	2,43	011	10,23	15
240307003	Géométrie	6,09	10	1,76	006	6,00	10
240307004	Plan cartésien	8,14	15	2,81	007	8,19	15
240307	Mathématique globale	31,62	50	6,70	030	31,12	40

2.2 Indice de réussite et de discrimination (analyse d'items)

Le tableau ci-dessous présente l'indice de succès (P) et de discrimination (Rbis) pour chacun des items du test. L'indice de discrimination est donné uniquement pour l'ensemble du Québec. On ne le retrouvera donc pas dans les statistiques fournies aux organismes. Cet indice permet de juger de la valeur des items. Pour une meilleure interprétation du tableau, on pourra se rapporter au manuel d'information du test qui donne la liste des objectifs de chacun des items, ou aux cahiers du test eux-mêmes.

TABLEAU 2 - Test de mathématique 240 307
Indice de réussite (P) et de discrimination (Rbis)
Echantillons de 2500 de l'ensemble de la population et de chacune des voies 281-422 et 281-432.

AVRIL 1979

NO DE L'ITEM	% DE SUCCES (P) Ensemble de la population	VOIE 281-422	VOIE 281-432	INDICE DE DISCRIMINATION (Rbis) *
001 - Ensembles et systèmes de nombres (10 items)				
1	76,0	63,9	81,4	0,61
2	66,7	50,4	81,5	0,61
3	58,7	59,9	70,5	0,50
4	57,6	47,8	65,8	0,65
5	73,5	64,5	84,5	0,58
6	53,6	38,6	61,4	0,53
7	48,4	39,8	55,3	0,46
8	35,3	28,2	51,0	0,38
9	61,4	58,7	70,1	0,34
10	36,4	30,0	45,6	0,40
Moyenne au sous-test	56,5	49,0	66,5	
002 - Calcul algébrique et équations (15 items)				
11	77,9	60,4	88,7	0,46
12	70,7	65,0	82,3	0,64

TABLEAU 2 (suite)

NO DE L'ITEM	% DE SUCCES (P) Ensemble de la population	201-422	201-452	INDICE DE DISCRIMINATION (Rbis) *
002 (suite)				
13	61,0	61,0	76,1	0,76
14	66,8	52,0	76,4	0,60
15	56,9	47,4	76,8	0,64
16	57,5	45,9	68,6	0,72
17	51,9	45,6	65,6	0,61
18	60,8	51,9	76,6	0,72
19	64,6	51,9	77,7	0,57
20	42,1	34,0	51,4	0,52
21	45,1	28,8	64,4	0,71
22	46,8	39,2	52,1	0,48
23	39,8	28,0	49,9	0,60
24	38,6	26,3	45,1	0,60
25	47,5	31,4	53,8	0,65
Moyenne au sous-test	55,3	44,4	70,3	
003 - Géométrie (10 items)				
26	75,4	73,2	80,4	0,27
27	75,7	72,7	89,0	0,55

TABLEAU 2 (suite)

NO DE L'ITEM	% DE SUCCES (P) Ensemble de la population	281-422	281-432	INDICE DE DISCRIMINATION (Rbis) *
003 (suite)				
28	75,4	72,7	80,1	0,36
29	63,8	59,9	72,0	0,52
30	55,5	53,0	64,9	0,48
31	41,2	38,5	57,2	0,28
32	30,8	26,4	48,2	0,49
33	52,7	37,4	67,4	0,74
34	37,8	36,8	56,7	0,55
35	15,0	9,6	17,4	0,34
Moyenne au sous-test	52,3	48,0	67,1	
004 - Fonctions et plan cartésien (15 items)				
36	59,6	54,8	71,9	0,57
37	54,0	48,1	71,4	0,70
38	66,5	57,8	77,6	0,64
39	66,3	57,9	76,3	0,66
40	48,6	44,7	63,8	0,52
41	53,4	44,0	73,3	0,68

TABLEAU 2 (suite)

NO DE L'ITEM	% DE SUCCES (P) Ensemble de la population	201-422	201-432	INDICE DE DISCRIMINATION (Rbis) *
004 (suite)				
42	42,2	24,8	55,1	0,64
43	52,1	45,2	66,2	0,47
44	34,0	31,3	40,5	0,46
45	36,4	30,2	44,0	0,55
46	29,0	19,2	40,2	0,46
47	36,7	24,7	42,7	0,48
48	34,4	26,8	49,6	0,47
49	32,2	26,4	41,9 ?	9,35
50	32,4	19,5	40,8 ?	0,56
Moyenne au sous-test	45,2	37,0	58,1	
Moyenne à l'ens. du test	51,9	42,0	63,2	

* item vs test entier

2.2.1 Interprétation - Indice de réussite et de discrimination

(A) L'indice de réussite (P)

Le tableau 2 indique le degré de réussite de chacun des items exprimé en pourcentage de succès (P), en considérant un échantillon au hasard (N) de 2500 candidats pour tout le Québec. Cet indice de réussite est donné pour un échantillon global, de même que pour un échantillon de candidats inscrits selon les voies 281-422 et 281-432.

Peut-on s'attendre à un succès de 50%, 60% ou 85% si l'on admet la pertinence des items étudiés? Bien que l'évaluation d'une performance particulière soit d'abord une question locale, il peut être utile de comparer les résultats d'un groupe à ceux de toute la population.

Cependant les données obtenues pour la population globale ne peuvent servir de critères de maîtrise. Ce n'est pas parce qu'un item est réussi à 90% au niveau du Québec que l'on doit considérer cette performance comme l'idéal à atteindre quand il s'agit de déterminer la norme de maîtrise pour l'aspect mathématique mesuré par cet item. De même ce n'est pas parce qu'un item est réussi à 10% que l'on peut dire que l'aspect mathématique mesuré par cet item dépasse le niveau de compréhension des élèves et devrait être supprimé des programmes. Tout ce qu'on peut dire, c'est que tel aspect mesuré, sous telle forme, à tel moment donné, est réussi à tel ou tel pourcentage. Le test nous renseigne sur ce qui est et non sur ce qui devrait être. Aucun test ne nous dit ce qui doit être enseigné, comme tel, mais il peut bien nous renseigner sur les acquisitions ou les ignorances des élèves suite à l'enseignement. Il décrit une situation et les mêmes interrogations se posent au niveau de l'ensemble de la population comme au niveau local. Il faudra pousser la recherche plus loin, à partir des indices que le résultat peut fournir, pour en arriver à des conclusions plus précises.

(B) L'indice de discrimination (Rbis)

Le tableau 2 indique également le degré de discrimination de chacun des items pour un échantillon de candidats, sans distinction d'inscription.

Dans un examen ou un test bien construit, une question que l'on juge "bonne" devrait être réussie par les meilleurs élèves dans le domaine mesuré, c'est-à-dire ceux qui réussissent le mieux à l'ensemble de l'examen ou du test. L'indice de discrimination utilisé ici se pose de nous dire si telle ou telle question est "bonne"; il met en relation le succès à chacune des questions et le succès à l'ensemble du test. Il nous dit si les deux vont dans le même sens. L'indice de discrimination utilisé est le coefficient de corrélation bisérial (Rbis).

Cet indice varie de -1 à +1. Plus il est élevé (près de +1), plus nous avons de chance de retrouver les élèves, qui ont réussi l'item dont on veut juger de la valeur, dans le groupe de ceux qui ont le mieux réussi à l'ensemble du test. Un indice de 0 nous indique qu'il n'y a aucune corrélation entre deux variables, tandis qu'un indice de -1 indique une corrélation négative (1).

On considère généralement comme significatif un Rbis dont l'indice est +0,30 et plus. Cependant certains items dont l'indice peut être légèrement inférieur à +0,30, mais quand même positif, ont été gardé en raison de l'importance de l'objectif mesuré.

(1) On trouvera un exposé sur ce qu'il est essentiel de connaître au sujet de l'indice de discrimination dans le Guide docimologique, fascicule no 6, publié par le MEQ (disponible à la DMEA).

2.3 Répartition des résultats

Les figures 2 à 6 présentent la distribution des résultats obtenus à chacun des sous-tests et au test entier. Elles permettent de situer les différents scores et le pourcentage de candidats (P%) qui ont obtenu ces scores au niveau du Québec. On pourra, au niveau local, établir des comparaisons avec les histogrammes de fréquence pour les différents aspects du test.

Ces figures établissent également des comparaisons avec les résultats obtenus au test et aux sous-tests pour l'année 1977-78.

Figure 2 à 6 - Courbes de distribution des scores bruts
Avril 1978 et 1979

Figure 2 - Sous-test 001 - Nombres et symboles
de nombres
Moyenne (M): 5,59

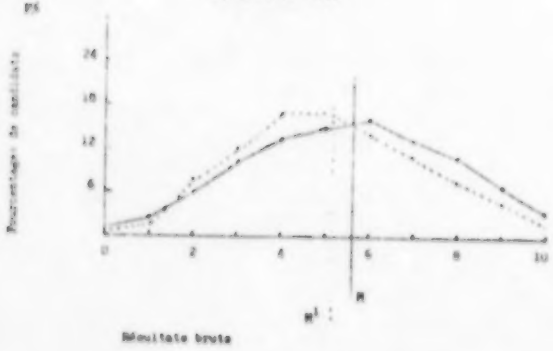


Figure 3 - Distribution des scores bruts
Sous-test 002 - Calcul algébrique et équations
Moyenne (M): 8,19

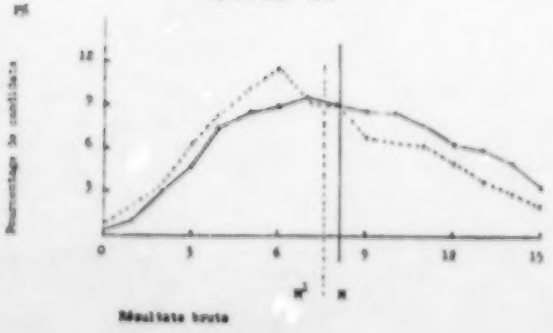


Figure 4 - Distribution des scores bruts
Sous-test 105 - géométrie
Moyenne (M) : 5,17

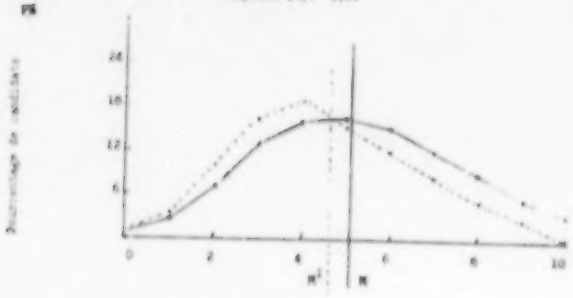


Figure 5 - Distribution des scores bruts
Sous-test 104 - Fonctions - Plan cartésien
Moyenne (M) : 6,67

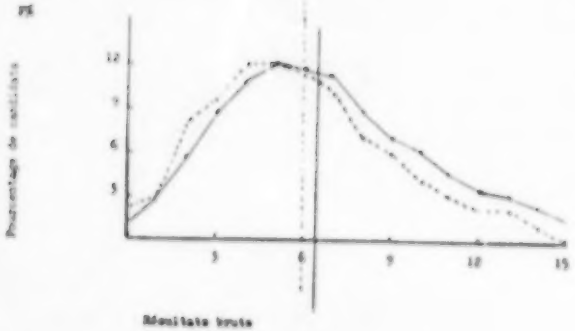
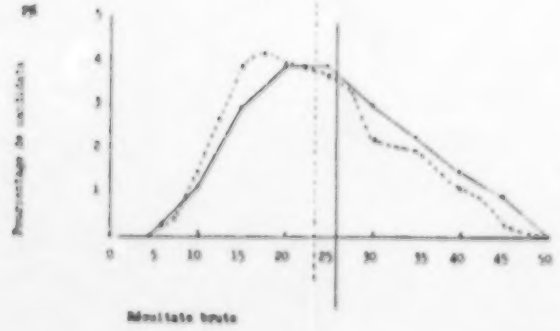


Figure 6 - Distribution des scores bruts
Ensemble du test
Moyenne (M) : 25,63



——— administration avril 1979
- - - - administration avril 1978

2.4 La liste des normes

Les tableaux 3A et 3B ci-dessous décrivent, pour l'ensemble des tests et pour chacun des sous-tests, la performance que peut représenter chacun des résultats bruts obtenus par la population de l'ensemble du Québec, qui a passé le test en avril 1979 (23 913). Cette performance est décrite en côtes standardisées, en stanines et en centiles. Pour une meilleure compréhension des normes, on pourra consulter le tableau d'équivalence (à la page). A la lumière de ces données les organismes pourront interpréter les normes locales.

Le tableau 4 décrit à titre indicatif la performance que peut représenter les résultats bruts globaux obtenus au test, si les candidats étaient classés séparement selon les différentes voies 281-422 et 281-432. Ce tableau permet une meilleure interprétation des normes pour l'ensemble de la population.

Si l'on désire interpréter les résultats individuels en fonction du tableau 4, on pourra se servir des résultats bruts fournis dans la liste des résultats fournis aux organismes. Les tableaux 1B et 1C (p.17) permettront une meilleure interprétation du tableau 4.

TABLEAU 4A - Test de mathématique 240 507 - résultat global
Normes - ensemble du Québec, avril 1979
Population totale (23 915)

RÉSULTAT BRUT	STANINE	RANG CENTILE	SCORE STANDARDISÉ	RÉSULTAT BRUT	STANINE	RANG CENTILE	SCORE STANDARDISÉ
0	1	0	+34	26	5	52	+63
1	1	0	+35	27	5	56	+64
2	1	0	+36	28	5	59	+65
3	1	0	+38	29	6	63	+66
4	1	0	+39	30	6	66	+67
5	1	0	+40	31	6	69	+68
6	1	0	+41	32	6	72	+70
7	1	0	+42	33	6	75	+71
8	1	0	+43	34	7	78	+72
9	1	1	+44	35	7	80	+73
10	1	2	+45	36	7	82	+74
11	1	3	+46	37	7	85	+75
12	2	4	+47	38	7	87	+76
13	2	6	+49	39	7	88	+77
14	2	8	+50	40	8	90	+78
15	3	11	+51	41	8	92	+79
16	3	14	+52	42	8	93	+81
17	3	18	+53	43	8	94	+82
18	3	21	+54	44	8	96	+83
19	4	25	+55	45	9	97	+84
20	4	29	+56	46	9	97	+85
21	4	33	+57	47	9	98	+86
22	4	37	+58	48	9	99	+84
23	5	41	+60	49	9	99	+88
24	5	45	+61	50	9	99	+89
25	5	48	+62				

* le chiffre 0 indique l'insuffisance des données.

TABLEAU 3B - Test de mathématique 240 307 - sous-tests
Normes - ensemble du Québec, avril 1979
Population totale (25 913)

RÉSULTAT BRUT	STANINE	RANG CENTILE	SCORE STANDARDISÉ	STANINE	RANG CENTILE	SCORE STANDARDISÉ	RÉSULTAT BRUT
001 - Ensemble et système de nombres				003 - Géométrie			
0	1	0	+37	1	0	+38	0
1	1	0	+42	1	0	+43	1
2	1	3	+46	1	3	+48	2
3	2	8	+51	2	10	+52	3
4	3	19	+55	4	23	+57	4
5	4	32	+60	4	39	+62	5
6	6	48	+64	5	56	+66	6
7	7	64	+69	6	71	+71	7
8	8	77	+73	7	84	+76	8
9	9	89	+78	8	93	+81	9
10	10	96	+82	9	98	+85	10
002 - Calcul algébrique et équations				004 - Fonctions - Plan cartésien			
0	1	0	+39	1	0	+43	0
1	1	0	+42	1	1	+46	1
2	1	1	+45	2	4	+49	2
3	2	4	+48	2	9	+52	3
4	2	10	+51	3	18	+55	4
5	3	17	+54	4	29	+58	5
6	4	26	+56	5	41	+61	6
7	4	35	+59	5	52	+65	7
8	5	45	+62	6	62	+66	8
9	5	54	+65	6	71	+69	9
10	6	62	+68	7	78	+72	10
11	6	70	+70	7	84	+75	11
12	7	78	+73	7	89	+78	12
13	7	85	+76	8	93	+81	13
14	8	91	+79	8	96	+84	14
15	8	96	+86	9	98	+87	15

TABLEAU 4 - Test de mathématique 240 307 - résultat global
Normes - ensemble du Québec à partir d'un échantillon de 2 500 candidats choisis au hasard
Voies 281-422 et 281-432

Math 240 307 - Normes - ensemble du Québec - 782 Voie (281-422) 783 Voie (281-432)

RÉSULTAT BRUT	SCORES STANDARDISÉS 281-422	281-432	RÉSULTAT BRUT	SCORES STANDARDISÉS 281-422	281-432
0	28	16	26	71	55
1	30	17	27	73	57
2	33	19	28	75	58
3	34	20	29	76	59
4	35	22	30	78	60
5	36	23	31	80	62
6	38	25	32	82	63
7	40	26	33	83	65
8	41	28	34	85	66
9	43	29	35	86	68
10	44	30	36	88	70
11	46	32	37	89	72
12	48	33	38	90	73
13	49	35	39	92	74
14	51	37	40	93	75
15	53	38	41	95	76
16	54	40	42	96	78
17	56	41	43	98	79
18	58	43	44	99	81
19	60	44	45	101	83
20	62	45	46	103	85
21	63	47	47	105	88
22	64	48	48	106	90
23	66	50	49	108	91
24	68	51	50	110	93
25	70	53			

2.4.1 L'utilité des normes

Le résultat brut est toujours pertinent à une partie du test ou au test global. Il ne peut être comparé directement avec le score obtenu à un autre sous-test. Deux résultats identiques obtenus à deux différents sous-tests peuvent représenter des niveaux de performance différents, compte tenu de certains facteurs comme la longueur des sous-tests, leur difficulté relative, etc. On doit admettre que les scores bruts n'ont aucune signification en eux-mêmes, à moins que l'on puisse établir des standards de référence qui permettent de les jauger. Mais il est difficile de juger de la difficulté du travail assigné aux candidats dans l'une ou l'autre des parties du test.

Une bonne méthode consiste alors à contourner la difficulté en établissant des normes, c'est-à-dire partir de la performance des autres candidats. Selon les références adoptées et la précision des données, les normes nous permettent de dire que tel ou tel résultat brut représente telle ou telle performance en rapport à un groupe particulier de candidats: école, organisme, région, ensemble du Québec.

Dans le cas où un résultat brut se distribue sur plusieurs centiles, on indique le centile le plus bas. Ainsi dans le tableau 4, le score brut 20 indique que les candidats qui ont obtenu ce score se distribuent entre les centiles 39 à 42; on indique alors dans le tableau que le centile 39 correspond à ce score. La même remarque s'applique aux scores standardisés.

3. CONSIDÉRATIONS TECHNIQUES

On trouvera dans les pages qui suivent certaines données techniques. Celles-ci peuvent être utiles dans une interprétation plus judicieuse des résultats et de la valeur du test.

3.1 La validité du test

On doit se demander qu'elle est la question la plus importante: "Est-ce que le test permet de bien différencier les élèves?" (test normatif) ou plutôt "Est-ce que le test ne devrait pas mesurer ce que les élèves peuvent savoir?" (test critérié). Si l'on accorde plus d'importance à l'aspect considéré dans la dernière question, il faut, dans ce cas, établir d'abord la validité de contenu.

On reconnaît généralement que cette dernière est essentielle pour les instruments qui mesurent le rendement scolaire, catégorie dans laquelle on peut classer les tests d'évaluation de la qualité de l'apprentissage (Assessment tests).

3.1.1 La validité de contenu: la pertinence du test vis-à-vis les objectifs d'apprentissage

Pour apprécier cette validité, il faudra voir jusqu'à quel point le test représente un échantillon de la matière à maîtriser et des comportements qui doivent être acquis; l'élaboration du test peut représenter ainsi un problème d'échantillonnage vis-à-vis les contenus mathématiques et les processus cognitifs à maîtriser. Il faut donner à chaque aspect à mesurer l'importance relative que celui-ci doit avoir. De plus, il faudra non seulement considérer le contenu des questions, mais les processus présumément utilisés par les élèves pour obtenir la réponse aux différentes questions du test; quel comportement est mis en cause dans le type de question utilisées?

Elaboration du contenu

Pour atteindre cette validité dans le test de mathématique, les auteurs ont d'abord dressé une liste des objectifs d'apprentissage qui leur paraissaient les plus pertinents à la fin de la quatrième secondaire. Ces objectifs de caractère plutôt cognitifs devraient être atteints par la plupart des élèves, quel que soit l'enseignement offert. Il se peut qu'un objectif présente un caractère global et vise plus d'un comportement, mais dans les limites du programme-cadre du MEQ et des programmes institutionnels déjà élaborés depuis plusieurs années.

A partir de ces objectifs dit "terminaux", des spécialistes du domaine mesuré ont établi les spécifications du test, rédigé et revisé les items en fonctions de ces spécifications. On comprendra qu'un tel test (50 items) est forcément limité, qu'il ne peut couvrir tous les objectifs d'apprentissage même importants. La mesure de certains aspects est plus particulièrement limitée. En géométrie, entre autres, on s'est borné à certains aspects qu'on peut juger très élémentaires et qui ne dépassent pas le niveau taxonomique de connaissance. De ce point de vue, le test constitue un premier sondage. Les auteurs ont du s'assurer au moins que tous les aspects mesurés sont pertinents. Une pré-expérimentation a permis de faire un rajustement et de rejeter les items non valides.

Cependant, dans l'interprétation, on devra tenir compte que l'apprentissage se déroule dans le temps. Le test peut mesurer des apprentissages à un moment où il reste encore quelques mois d'apprentissage; l'acquisition de certains aspects dans certains programmes institutionnels peut être prévue au cours de ces quelques mois. Il en résulte une variation des résultats. L'analyse du contenu du test, de ce point de vue, permettra de tenir compte de ce fait dans l'interprétation.

Toutefois, il y a tout lieu de croire que le contenu du test de mathématique est en relation logique avec ce qui en constitue "le rationnel".

3.1.2 Standardisation et "erreur d'échantillonnage"

Les normes d'un test sont généralement établies à partir des résultats obtenus par un groupe de candidats représentatif de la population totale de ceux à qui le test s'adresse. On n'est cependant jamais sûr de la représentativité du groupe vis-à-vis de la population totale; ceci entraîne une certaine erreur dite "d'échantillonnage" dans les normes. Tel n'est pas le cas des tests du MEQ: ceux-ci étant passés à toute la population ou presque, leurs résultats normalisés y réfèrent directement.

3.1.3 Corrélation entre les sous-tests et le test entier

Le tableau 5 permet de juger des relations entre chacun des sous-tests et le test entier. Il donne des coefficients de corrélation de Pearson.

Les corrélations entre les sous-tests, comme on le voit, sont relativement élevées. Ceci suggère que les connaissances ou les habiletés mesurées par chacun des sous-tests sont en relation assez étroite avec les autres parties, ce qui est tout à fait normal si l'on considère les rapports entre les différents aspects du contenu mathématique.

TABLEAU 5 - Test de mathématique 240 307
Corrélations entre les sous-tests et le test entier - avril 1979

SOUS-TEST OU TEST	001	002	003	004	005
001 - Ensembles et systèmes de nombres					
002 - Calcul algébrique et équations	0,638				
003 - Géométrie	0,484	0,543			
004 - Fonctions et plan cartésien	0,519	0,634	0,525		
005 - Test entier	0,780	0,886	0,740	0,849	

3.2 La fidélité

3.2.1 La consistance interne

La fidélité d'un test peut se vérifier par la constance des informations que le test peut fournir. Dans ce cas, il s'agit d'établir si le test appliqué plusieurs fois dans les mêmes conditions fournit des résultats identiques.

Ainsi il s'agit d'établir la cohérence et la stabilité du test. Elle se vérifie par une mesure de consistance interne et elle est généralement exprimée par un coefficient de correlation.

Procédé utilisé

La consistance interne du test a été éprouvée par le calcul du Kuder-Richarson, formule 20 (KR-20), pour l'ensemble du test, et pour chacun des sous-tests. Ces données statistiques sont contenues dans le tableau 6.

Si les auteurs ont d'abord recherché une validité de contenu, ce coefficient nous renseigne toutefois sur son homogénéité, par la mise en relations des items. Si le KR-20 a une valeur de 0,0, ceci indiquerait qu'il n'y a aucune correlation entre les items, tandis qu'une valeur de +1,0 indiquerait une correlation positive parfaite.

Ici le KR-20 a une valeur positive de 0,89 pour l'ensemble du test. Si l'on considère que le test est relativement court et que les conditions d'administration peuvent varier d'un milieu à un autre, on peut dire que ceci est excellent et que le test mesure avec consistance ce qu'il devrait mesurer. Des batteries de mesure beaucoup plus considérables acceptent comme satisfaisantes des valeurs qui ne sont pas beaucoup plus élevées. "La plupart des tests que les auteurs considèrent excellents acceptent des coefficients de fidélité entre 0,60 et 0,80"(1).

Dans le cas des aspects mesurés, on tiendra compte du nombre d'items qui influence grandement la valeur du KR-20; ainsi dans les aspects 001 (ensembles et systèmes de nombres) et 003 (géométrie) on obtient respectivement 0,61 et 0,60 mais chacun des aspects ne compte que 10 items.

(1) Diederich, Paul B., Short-Cut Statistics for Teachers-Made Tests, E.T.S. Princeton N.J., (1973), p. 10.

De plus, on considère qu'il ne s'agit pas ici d'un test normatif de classement ou de placement, où les exigences pourraient être beaucoup plus grandes - dans ce cas, on pourrait s'attendre à au moins 0,90 - mais d'un test qui vise la mesure de la qualité de l'apprentissage.

3.2.2 L'erreur de mesure

Dans l'interprétation des résultats individuels, l'erreur de mesure est plus utile que le coefficient de consistance interne.

Parce que toute mesure est sujette à erreur, les résultats individuels obtenus au test se rapprochent plus ou moins des "résultats vrais", c'est-à-dire de ceux qu'on obtiendrait si on pouvait éliminer toutes les sources d'erreur qui entraîne des fluctuations. En pratique, l'erreur de mesure nous indique que le résultat obtenu ne doit pas être considéré d'une façon absolue et elle nous fournit une façon raisonnable de l'interpréter.

Ainsi en supposant qu'un candidat a obtenu 30 à l'ensemble du test, alors que l'erreur de mesure est de 3, selon le tableau, on dira que le candidat a 2 chances sur 3, selon des calculs statistiques éprouvés, d'obtenir 30 ± 3, au test, soit entre 27,0 et 33,0 comme résultat "vrai".

TABLEAU 6 — TEST DE MATHÉMATIQUE 243 317

CONSISTANCE INTERNE ET AUTRES INDICES RELATIFS - AVRIL 73

TEST	NOMBRE D'ITEMS	N	KR-20	MOYENNE	ÉTENDUE DES SCORES	ÉCART TYPE	VARIANCE	ERREUR DE MESURE
001 - ensembles et systèmes de nombres	10	2500	0,61	5,65	0 - 10	2,24	5,01	1,40
002 - calcul algébrique et équations	15	2500	0,78	8,29	0 - 15	3,57	12,75	1,66
003 - géométrie	10	2500	0,60	5,23	0 - 10	2,15	4,65	1,51
004 - fonctions et plan cartésien	15	2500	0,75	6,78	0 - 15	3,44	11,84	1,71
009 - résultat global	50	2500	0,89	25,96	4 - 50	7,41	54,15	3,12

Tableau 7 - Correlations calculées à partir du R de Pearson entre les sous-tests, le global du test de mathématique 240 307 (var. 1 à 5) et certains examens de rendement scolaire en secondaire 4 - mai 1979.

	VAR001	VAR002	VAR003	VAR004	VAR005
BIO412	0.5108 (40) P=0.000	0.6709 (40) P=0.000	0.3699 (40) P=0.009	0.4606 (40) P=0.001	0.6364 (40) P=0.000
EDPHYG	0.2578 (302) P=0.000	0.2755 (301) P=0.000	0.1404 (302) P=0.007	0.2086 (300) P=0.000	0.2744 (302) P=0.000
EDPHYF	-0.1026 (39) P=0.267	0.1241 (39) P=0.226	0.1484 (39) P=0.184	0.2868 (39) P=0.038	0.1674 (39) P=0.154
CHI442	0.2174 (26) P=0.143	0.3874 (26) P=0.025	0.7277 (26) P=0.000	0.3566 (23) P=0.047	0.5095 (26) P=0.004
MAT412	0.4511 (368) P=0.000	0.4625 (367) P=0.000	0.4122 (368) P=0.000	0.4675 (365) P=0.000	0.5471 (368) P=0.000
MAT422	0.3839 (194) P=0.000	0.5140 (194) P=0.000	0.2485 (194) P=0.000	0.3272 (191) P=0.000	0.5716 (194) P=0.000
MAT432	0.6158 (193) P=0.000	0.6982 (192) P=0.000	0.5287 (193) P=0.000	0.6491 (193) P=0.000	0.8043 (193) P=0.000
PHY422	0.5363 (246) P=0.000	0.5650 (245) P=0.000	0.5447 (246) P=0.000	0.4894 (246) P=0.000	0.6662 (246) P=0.000
PHY432	0.4337 (54) P=0.001	0.4768 (54) P=0.000	0.2189 (54) P=0.056	0.4413 (54) P=0.008	0.6327 (54) P=0.000
F422CT	0.4528 (238) P=0.000	0.4865 (238) P=0.000	0.3442 (238) P=0.000	0.4989 (237) P=0.000	0.6486 (238) P=0.000
F432CT	0.4507 (111) P=0.000	0.5488 (110) P=0.000	0.4091 (111) P=0.000	0.4603 (110) P=0.000	0.5795 (111) P=0.000

(Coefficient / population / degré de signification)

Le tableau 7 fournit les résultats partiels d'une étude préliminaire qui a été entreprise dans le but de mesurer la validité concomitante du test de mathématique en secondaire 4. Elle établit les relations entre certaines mesures du rendement scolaire: Biologie 412, éducation physique - garçons et filles, chimie 412, histoire 412, mathématique 422 et 432, physique 422 et 452, français 422 et 432 (compréhension de texte), et d'autre part les sous-tests (variables 1 à 4) ainsi que le global (variable 5) du test de mathématique 240 307.

Cependant il faut tenir compte des limites de ces comparaisons, qui risquent fort d'être biaisées.

Les corrélations entre deux mesures relatives sont sujettes aux limites des mesures appliquées. Plus la mesure est étendue, plus les résultats sont fidèles et par conséquent l'effet du hasard est diminué. Or si des corrélations entre l'ensemble des items d'un test (KR-20) ne dépassent pas .60 (consistance interne), il devient illusoire d'attendre plus de .60 entre cette mesure et un critère externe.

Les examens qui forment la deuxième série de variables dans le présent tableau de corrélations constituent des mesures relatives. La consistance interne de ces examens n'a pas été établie, mais on peut supposer qu'elle ne dépasse pas .60 étant donné le nombre de questions sur lesquelles se distribuent les résultats. Les objectifs principaux de ces examens n'étant pas de discriminer, mais plutôt de mesurer un degré minimal de compétence, plusieurs items sont réussis par tous, alors que tous les élèves échouent à d'autres et par conséquent, l'étendue des distributions peut être restreinte. On doit considérer également les échantillons restreint qui ont servi à établir ces comparaisons dans certains cas.

Tenant compte des explications qui précèdent, il n'est certes pas facile d'obtenir un coefficient de corrélation supérieur à .50. Le présent tableau renseigne toutefois sur la tendance des corrélations entre le test et des examens portant sur les contenus de certaines matières.

CONCLUSION - QUI INTERPRÈTE LES RÉSULTATS DES TESTS?

Bien que le MEQ publie les résultats pour les individus et les écoles ainsi que les données nécessaires qui aident à leur interprétation, ce n'est pas son rôle de les évaluer. Trop de circonstances lui échappent pour en mesurer la portée au niveau des candidats, des écoles et même des régions.

Toutefois, les données d'ensemble pour le Québec pourront fournir au MEQ des renseignements utiles pour l'évaluation et l'orientation des programmes d'études.

C'est en définitive aux personnes concernées, parents, enseignants, administrateurs que reviennent l'interprétation et l'évaluation des résultats du test. Ces personnes sont les mieux placées pour tenir compte de toutes les circonstances qui influent sur la performance des élèves.

Il est évidemment nécessaire que toutes ces personnes utilisent les résultats avec discernement et tiennent compte des circonstances qui en limitent la portée.

En tout temps, il est important de rappeler que les résultats aux tests standardisés doivent être interprétés avec prudence. Le contenu des tests limité face aux programmes d'apprentissage, leur court temps de passation et les circonstances particulières qui entourent toute mesure ainsi structurée peuvent révéler des faiblesses artificielles chez certains élèves.

Pour obtenir des renseignements supplémentaires, veuillez vous adresser à:

Service des tests
Direction de la mesure et de l'évaluation
des apprentissages
Ministère de l'Education
1035 rue de la Chevrotière
Québec
G1R 5A5
Tél.: (418) 643-1661

Coordonnateur du service: Monsieur Jacques Emond

Responsable des tests de mathématique: Monsieur Jean-Marcel Bourgault